DOCUMENT RESUME

ED 314 747 CS 212 184

AUTHOR Gentner, Dedre
TITLE Metaphor as Structure Mapping: The Relational Shift. Technical Report No. 488.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Jan 90
CONTRACT G0087-C1001-90
NOTE 27p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Adults; Cognitive Development; Cognitive Processes; *Cognitive Structures; Comparative Analysis; Concept Formation; Developmental Psychology; *Interpretive Skills; *Metaphors; *Young Children
IDENTIFIERS Metaphorical Thought; *Structure Mapping Theory

ABSTRACT

A study used structure mapping theory to examine the claim that children can understand metaphors based on shared object attributes before understanding those based on shared relational structure. The three kinds of metaphor used in the study varied in whether the shared information forming the basis for the interpretation was attributional, relational, or both. In the first experiment, 10 five-year-olds, 10 nine-year-olds, and 10 college students enrolled in psychology classes, produced interpretations of the three types of metaphors. The attributionality and relationality of their interpretations were scored by independent judges. In the second experiment, 36 adult undergraduates, 18 four-year-olds, and 18 seven-year-olds chose which of two interpretations (relational or attributional) of metaphor they preferred. In both experiments, results indicated that relational responding increased significantly with age, but attributional responding did not. Findings suggest a developmental shift toward a focus on relational structure in metaphor interpretation. (Three tables of data, three figures, and five footnotes are included; 39 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED314747

CENTER FOR THE STUDY OF READING

Technical Report No. 488

METAPHOR AS STRUCTURE MAPPING: THE RELATIONAL SHIFT

Dedre Gentner

University of Illinois at Urbana-Champaign

January 1990

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

CS212184

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. G0087-C1001-90 with the Reading Research and Education Center. The publication does not necessarily reflect the views of the agency supporting the research.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
D. Anderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EDITORIAL ADVISORY BOARD
1989-90

James Armstrong
Linda Asmussen
Gerald Arnold
Yahaya Bello
Diane Bottomley
Catherine Burnham
Candace Clark
Michelle Commeyras
John M. Consalvi
Christopher Currie
Irene-Anna Diakidoy
Barbara Hancin
Michael J. Jacobson
Jihn-Chang Jehng
Robert T. Jimenez
Bonnie M. Kerr
Paul W. Kerr
Juan Moran
Keisuke Ohtsuka
Kathy Meyer Reimer
Hua Shu
Anne Stallman
Marty Waggoner
Janelle Weinzierl
Pamela Winsor
Marsha Wise

MANAGING EDITOR
Fran Lehr

MANUSCRIPT PRODUCTION ASSISTANTS
Delores Plowman
Debra Gough

Abstract

The goal of this research is to clarify the development of metaphor by using structure-mapping theory to make distinctions among kinds of metaphors. In particular, it is proposed that children can understand metaphors based on shared object attributes before those based on shared relational structure. This predicts (a) early ability to interpret metaphors based on shared attributes, (b) a developmental increase in ability to interpret metaphors based on shared relational structure, and (c) a shift from primarily attributional to primarily relational interpretations for metaphors that can be understood in either way. Two experiments were performed to test these claims. There were three kinds of metaphors, varying in whether the shared information forming the basis for the interpretation was attributional, relational, or both. In Experiment 1, children aged 5-6 and 9-10 and adults produced interpretations of the three types of metaphors. The attributionality and relationality of their interpretations were scored by independent judges. In Experiment 2, children aged 4-5 and 7-8 and adults chose which of two interpretations--relational or attributional--of a metaphor they preferred. In both experiments, relational responding increased significantly with age, but attributional responding did not. These results indicate a developmental shift toward a focus on relational structure in metaphor interpretation.

METAPHOR AS STRUCTURE MAPPING: THE RELATIONAL SHIFT

The study of metaphoric development is fraught with contradictions. Evidence that young children are quite poor at metaphoric interpretation sits side by side with equally compelling evidence that they are uniquely talented at metaphoric language. In experimental studies of metaphor comprehension, young children typically perform quite poorly. A 4-year-old asked "Can a person be sweet?" answers literally (e.g., "Not unless he was made out of chocolate") (Asch & Nerlove, 1960). Young children do badly at matching sentences with metaphorically related pictures (Kogan, 1975) and at choosing appropriate metaphorical completions for sentences (Gardner, Kircher, Winner, & Perkins, 1975). Not until 14 years of age can children explain metaphors such as "the prison guard was a hard rock." These and many other experimental results seem to indicate that metaphorical ability develops very slowly. Indeed, until recently, the dominant position was that metaphor develops only *after* the child has acquired basic competence at literal language (e.g., Inhelder & Piaget, 1958).

This dour assessment contrasts sharply with the impression gained from children's spontaneous speech, which is so full of creative metaphors that some observers have viewed early childhood as a period of linguistic genius (Gardner, 1974). For example, a 15-month-old girl used "moon" to refer not only to the moon but to a half grapefruit and a hangnail (Bowerman, 1982). A 2-year-old boy I observed remarked that a crescent moon was "bent, like a banana," and on another occasion jumped into a pile of pillows and announced "leafs." It is unlikely that all such extensions could be accounted for as errors (Thomson & Chapman, 1975). This seems to leave us with the paradox that young children can produce metaphors on their own but cannot comprehend them in experimental studies. Part of the disparity may be due to the kinds of tasks used to assess comprehension. For example, children can demonstrate metaphoric ability earlier with enactment tasks than with verbal explanation tasks (Gentner, 1977a, 1977b; Pollio & Pickens, 1980; Vosniadou & Ortony, 1986). But even with age-appropriate methodology, young children perform far less well in comprehension tasks than their fluent production would lead us to expect (e.g., Gentner & Toupin, 1986).

I suggest that an essential step in sorting out the developmental picture is clarifying what we mean by "metaphor." The term covers a number of different kinds of comparisons, varying in their complexity and in the nature of the commonalities they convey (Lakoff & Johnson, 1980). In this report I use the structure-mapping theory of analogy to derive distinctions among kinds of metaphors (Gentner, 1980, 1983, 1986). To preview the conclusions, I will argue that children can produce and comprehend metaphors based on common object attributes before metaphors based on common relational structure.

Structure mapping. The basic intuition of structure-mapping theory is that an analogy is a mapping of knowledge from one domain (the base) into another (the target) which conveys that a system of relations that holds among the base objects also holds among the target objects. Thus an analogy is a way of noticing relational commonalities independently of the objects to which those relations apply. For example, Carnot in 1824 explained heat flow by analogy with a waterfall. The analogy conveys that just as a gradient from a high level to a low level will cause water to flow, given a path, so a gradient from a high temperature to a low temperature will cause heat to flow, given a heat path. This is a typical analogy in that the higher order relational structures are identical in base and target if the proper low-order correspondences among objects and functions are made (i.e., water ----> heat; level ----> temperature; and water path ----> heat path).

In interpreting an analogy, people seek to put the objects of the base in one-to-one correspondence with the objects of the target so as to obtain maximum structural match.[1] The corresponding objects in the base and target do not have to resemble each other at all; object correspondences are determined by roles in the matching relational structures. Central to the mapping process is the principle of *systematicity*: People prefer to map systems of predicates linked by higher order relations with inferential import rather than to map isolated predicates. The systematicity principle is a structural expression of our tacit preference for coherence and deductive power in interpreting analogy. Besides analogy, other kinds of similarity matches can be distinguished in this framework, according to whether the match is one of relational structure, object descriptions, or both. As noted above, *analogies* discard object descriptions and map relational structure. *Mere appearance* matches are the opposite: They map

aspects of object descriptions and discard relational structure. *Literal similarity* matches map both relational structure and object descriptions.

Metaphor. Now let us apply this framework to metaphor. Metaphors can be divided into four partially overlapping categories: attributional metaphors, relational metaphors, double metaphors, and a category of complex metaphors that resist analysis as one-one mappings and that I will not consider here (Gentner, Falkenhainer, & Skorstad, 1987). Attributional metaphors[2] (e.g., "Her arms were like twin swans") are *mere appearance* matches: they convey common object attributes. Here, the attributes "long," "thin," and "graceful" can be mapped from the base domain of swans to the target domain of her arms. Relational metaphors (e.g., Shakespeare's "Look, he's winding up the watch of his wit; by and by it will strike") can be analyzed as *analogies*: they convey that the base and target share common relational structure. Here, the intended commonalities have nothing to do with the object attributes of a watch--a glass face, metal cogs, and so on; instead, the metaphor conveys the common relational structure of a person setting a mechanism that will later produce seemingly spontaneous external effects. Finally, besides pure relational and attributional metaphors, there are also double metaphors that are mixtures of the two, such as Verbrugge and McCarrell's (1977) example, "Plant stems are drinking straws for thirsty trees." This metaphor conveys both the common attributes "long, thin, tubular" and the common relational structure "raises fluids from a lower to a higher place in order to benefit some creature."

Adults have been found to prefer relational metaphors over attributional metaphors and to focus primarily on relational commonalities in interpreting metaphors (Gentner, 1986; Gentner & Landers, 1985). In a prior study on which the present research is based, subjects produced interpretations of metaphors and rated their aptness, having first written out descriptions of the objects contained in the metaphors (Gentner, 1986). The results indicated a relational focus in metaphor interpretation in two ways. First, although the object descriptions contained both object attributes and relational information, only the relational information survived into the metaphor interpretations. For example, one subject described "cigarette" as follows: "chopped cured tobacco in a paper roll/with or without a filter at the end/held in the mouth/lit with a match and breathed through to draw smoke into the lungs/found widely among humans/known by some cultures to be damaging to the lungs/once considered beneficial to health." This description contains both relational and attributional information. But the same person, when interpreting the metaphor "Cigarettes are like time bombs," included only relational information: "They do their damage after some period of time during which no damage may be evident." The second indication of relational focus is that subjects considered metaphors more apt when they could find relational interpretations: Subjects' aptness ratings for metaphors correlated positively with the judged relationality of their interpretations. In contrast, attributional commonalities did not contribute to aptness: The correlation between aptness ratings and attributionality was nonsignificant and negatively trending. Adults thus demonstrate a relational focus both in interpreting and in judging metaphor.

Do children show this same relational focus? Although research on metaphorical development has not in general addressed this question, there is reason to suspect that they do not. Billow (1975) asked children to interpret two kinds of metaphors: "similarity metaphors," based on shared attributes (e.g., "Hair is spaghetti") and "proportional metaphors," based on a class of shared relations (e.g., "My head is an apple without any core"). Children 9-14 years old were much better at interpreting the similarity metaphors than the proportional metaphors. Further, when given similarity metaphors even 5-year-olds could produce sensible interpretations almost 30% of the time. (They were not given the proportional metaphors.) These results suggest that attribute matches may be developmentally prior to relational matches. If this is true, then many of the apparent paradoxes in the literature on metaphor development may be traceable to different developmental patterns for different kinds of metaphorical matches. The present research investigates this possibility.

Experiment 1

Experiment 1 was a straightforward extension of the method used in the Gentner (1986) study. Children and adults interpreted metaphors of different types and rated the aptness of the metaphors. Their interpretations were then scored by independent judges for relationality and attributionality (see Gentner & Stuart, 1983, for further details).[3] There were three metaphor types: (a) *attributional* metaphors (e.g., "Pancakes are nickels" [both are round and flat]); (b) *relational* metaphors (e.g., "A tire is a shoe"

[both are used by moving figures as points of contact with the ground]); and (c) *double* metaphors, with both attributes and relations in common (e.g., "Plant stems are drinking straws" [both are long and cylindrical; both are used to bring liquids from below to nourish a living thing]).

If adults follow the predicted relational strategy, their responses should show three characteristics. First, relational and double metaphors should be interpreted relationally. (This prediction does not apply to the attribute metaphors because they are designed not to have legitimate relational interpretations.) Second, adults should prefer metaphors that allow relational interpretations, so their aptness ratings should be higher for relational and double metaphors than for attribute metaphors. Finally, adults' aptness ratings for metaphors should be positively correlated with the relationality of their metaphor interpretations but not with the attributionality. To the extent that children fail to follow a relational interpretation strategy, their results will differ from this pattern.

The experimental procedure was aimed at minimizing the influence of other developmental changes besides metaphorical development. The first potential confound is differences in pragmatic task knowledge: Young children may not know when they are supposed to give metaphorical answers. To circumvent such differences, a series of amplifications was used to make the task clear. A second possible confound is age differences in domain knowledge. To minimize such differences, only domains likely to be familiar to young children were used. Finally, to rule out differences in prior exposure to specific items, all comparisons were novel; no idioms or conventional metaphors were used.

Method

Subjects. There were 10 subjects from each of three age groups: 5-6-year-olds (5 boys and 5 girls), 9-10-year-olds (4 boys and 6 girls), and college students from psychology classes at the University of California at San Diego (7 males and 3 females). Children were recruited from schools in Del Mar and La Jolla, California.

Materials and design. There were eight instances each of three types of metaphor, as shown in Table 1: (1) attributional metaphors (A metaphors), in which base and target shared many attributes but few relations; (2) relational metaphors (R metaphors), in which base and target shared relational structure but few attributes; and (3) double metaphors (D metaphors), in which base and target shared both relations and attributes. Thus, each subject interpreted 48 objects and 24 metaphors. The design was Age (3 levels) x Metaphor Type (R, A, or D).

[Insert Table 1 about here.]

Procedure. The task was administered to the adults in written form, in groups. Subjects were first asked to describe the separate objects that later appeared in the metaphors. The objects were presented in semirandom order, with pairs from the same metaphor separated. All subjects, including children, were able to demonstrate basic familiarity with all of the terms used. After completing the object descriptions, subjects were given the metaphors. They were asked to write their interpretations of what the metaphors were intended to convey and also to rate the aptness of each metaphor (i.e., how clever, interesting, or worthwhile it was) on a 1-5 scale. They also rated the metaphoricity of each comparison (i.e., the degree to which it expressed literal similarity as opposed to metaphorical relatedness). Metaphoricity ratings were not elicited from children and are not considered further here. Order of presentation was semirandom, with the constraint that metaphors of the same type were separated.

The task was administered to children orally and individually, and responses were tape-recorded. Children were first asked to describe the 48 objects. Then they were asked to interpret the metaphors. A graded series of amplifications was used to be sure that children understood the task. The technique was essentially one of restating metaphors as similes, based on Reynolds and Ortony's (1980) finding that young children perform better with similes ("X is like a Y") than with metaphors ("X is a Y"). The experimenter would ask, "Is a hummingbird a helicopter?" If the child responded literally (e.g., "No, a hummingbird is a bird."), the experimenter asked, "What does it mean if I *say* 'A hummingbird is a helicopter'?" If the child still responded literally, the experimenter asked, "What does it mean if I say 'A hummingbird is *like* a helicopter'?" If the child maintained a literal response after this third question, the literal response was recorded and the experimenter went on to the next metaphor. This technique

was quite successful, eliciting nonliteral responses in all but a few cases. Moreover, after a few such sequences, most children caught on that a nonliteral response was desired in this task and thereafter produced nonliteral responses immediately for each new metaphor. After interpreting the metaphor, children rated its aptness by pointing to one of five schematic faces, ranging from very sad (low aptness) to very happy (high aptness). They were told to point to the happy face if they thought the metaphor was interesting and to the sad face if they thought it was dull or boring.

Scoring. The purpose of the scoring was to provide an unbiased assessment of the commonalities selected by the subjects in their metaphor interpretations: in particular, whether these commonalities were relational or attributional. Groups of from two to four advanced undergraduates who had received training in the use of propositional notation rated each metaphor interpretation on two separate 5-point scales, an attributionality scale and a relationality scale. Attributionality and relationality were scored on separate scales to allow for the possibility that the same response might contain both kinds of information. For the attributionality rating, judges rated whether the response described objects in and of themselves (e.g., "X is yellow," or "X is a sphere"). Most adjectives received high attributionality ratings, as did many concrete nouns. For the relationality rating, the judges scored whether the response described relations between objects (or, in the case of higher order relations, between other relations in the domain). Transitive verbs (e.g., "X chases Y," or "X causes Y"), comparative adjectives (e.g., "X is longer than Y"), and prepositions (e.g., "X is inside Y") tended to receive high relationality ratings.[4] The specific content of the relations could vary widely, including spatial, causal, functional, and other kinds of relations.

Two further points must be made about the scoring. First, to minimize the effects of differences in length of responses, an interpretation was given a high rating on relationality if it included any clearly relational statement, and similarly for attributionality. This method is sensitive to the presence or absence of relational (or attributional) information but not to the number of different relations (or attributes) mentioned. Second, because of the partly subjective nature of the scoring task, it was important to remove potential sources of bias. Therefore the judges were not told the subjects' aptness and metaphoricity ratings, and only one of the judges was aware of the experimental hypotheses. To conceal the ages of the subjects, the metaphor interpretations were scored in random order. Interrater agreement ranged from 85% to 100% on different metaphors.

Results

There was a marked developmental increase in the use of relational information in the metaphor interpretations. Table 2 shows sample interpretations produced at different ages, as well as the ratings they received. The mean rating of the relationality of metaphor interpretations increases steadily with age for both the relational (R) and double (D) metaphors (Fig. 1, top). (This increase does not occur for the attributional (A) metaphors because they do not permit relational interpretation.) In contrast, the attributionality of the metaphor interpretations remains constant across age for all three classes of metaphor (Fig. 1, bottom). In short, there is a specific developmental increase in the relationality, but not the attributionality, of the metaphor interpretations.

[Insert Figure 1 about here.]

[Insert Table 2 about here.]

Two separate two-way 3 (Age) x 3 (Metaphor Type) analyses of variance were performed, one for the relationality ratings and one for the attributionality ratings. In the relationality analysis, there was a main effect of age, confirming the developmental increase in use of relational information in interpreting metaphors, $F(2,27) = 12.76, p < .01$. The interaction of Age x Metaphor Type was also significant, reflecting the fact that the age increase in relationality occurs only for the R and D metaphors, $F(4,54) = 5.48, p < .01$. In the attributionality analysis, neither the main effect of age nor the Age x Metaphor Type interaction was significant, confirming that there was no developmental increase in attributionality. In both the relationality and attributionality analyses, the main effect of metaphor type was strongly significant, $F(2,54) = 191.63, p < .001$, $F(2,54) = 265.06\, p < .001$, respectively. For all ages, the a priori categorization of stimuli held up well: on relationality, R metaphors were highest, followed by D and

then A metaphors; and on attributionality, A metaphors were highest, followed by D and then R metaphors.

The performance on D metaphors is of special interest. By design, the D metaphors could support either an attributional or a relational interpretation. The results show a relational bias in the two older groups but not in the youngest children. Planned comparisons between the relationality ratings and the attributionality ratings of the D metaphors showed that, for the 9-year-olds and adults, the mean relationality is significantly greater than the mean attributionality, $t(9) = 2.78, p < .05$, and $t(9) = 3.79$, $p < .05$, respectively. For the 5-6-year-olds, there is no significant difference, $t(9) = 1.93, p > .05$.

Aptness ratings. As predicted, adults preferred metaphors for which they could find relational interpretations. Adults' mean aptness ratings for R($M = 2.86$) and D($M = 2.95$) metaphors were considerably higher than for A metaphors ($M = 2.30$), $t(79) = 4.30, p < .001$, for D versus A, and $t(79) = 3.87, p < .001$, for R versus A metaphors. Children did not show this relational preference. For 5-6-year-olds, $M = 2.08$, 2.14, and 2.26, and for 9-10-year-olds, $M = 1.99$, 2.18, and 2.19 for R, D, and A metaphors, respectively. Children's aptness ratings did not differ significantly across the three types of metaphors. Another indication of relational focus in adults is that their aptness ratings were positively correlated with the relationality of their interpretations and negatively correlated with the attributionality, $r(22) = .55, p < .01$, and, $r(22) = -.42, p < .05$, respectively. Children did not show this pattern. Their aptness ratings showed no correlation with either relationality, $r(22) = .15$ for 5-year-olds and -.18 for 9-year-olds, or attributionality, $r(22) = .01$ for 5-year-olds and .29 for 9-year-olds. We must be cautious in evaluating the aptness data because the children may simply have lacked facility with the aptness scale. Still, the results suggest a developmental difference in the criteria for judging aptness in metaphor.

Discussion

As predicted by structure-mapping theory, adults focused on mapping relational systems. There are several indications of this pattern. First, adult interpretations for R and D metaphors were highly relational; this is especially telling for D metaphors, which could have supported either kind of interpretation. Second, adults rated the R and D metaphors as more apt than the A metaphors. Third, adults' aptness ratings correlated positively with the relationality of their interpretations but negatively with attributionality. This pattern replicates and strengthens the positive correlation between aptness and relationality found in prior research with adults (Gentner, 1986). Adults appear both to seek relational predicates in metaphorical mapping and to judge the aptness of a comparison according to the relationality of the mapping.

Young children did not share this relational focus. The 5-6-year-olds produced fewer relational interpretations of R and D metaphors than the older groups, and they showed no tendency to produce relational rather than attributional interpretations for the D metaphors. There was a marked increase in relationality with age. This developmental trend was specific to relational information; there was no age increase in attributionality. Turning to aptness, one might have suspected that children, though not fluent enough to produce relational interpretations, would nevertheless show an appreciation of relationality in their aptness ratings. But this was not the case. Children's aptness ratings reflected no preference for R and D metaphors over A metaphors. Further, there was no correlation between children's aptness ratings for metaphors and the relationality of their own interpretations.

Overall, these results appear to show a clear developmental increase in relational focus. But there were some possible problems with the methodology used in Experiment 1. One limitation was the use of verbal explanations as responses. Children's ability to give an explicit verbal account of their knowledge typically lags behind their ability to demonstrate that knowledge in other ways (e.g., Pollio & Pickens, 1980; Vosniadou & Ortony, 1983, 1986). A second concern was the aptness-rating task. As discussed above, it was not clear that children understood the sad-to-smiling face scale. To address these problems, a second experiment was performed using a choice task rather than a production task. Children were presented with two interpretation--one relational and one attributional--for each metaphor and were asked to choose the interpretation they preferred. In this way, young children could express a preference for relational interpretations without having to generate them. After this choice task, a rating task was given, using a new method to elicit the ratings. Children were asked to indicate their

ratings on a "goodness meter," a vertical scale something like a thermometer. Finally, one other difference between the two studies was that, because Experiment 2 used a comprehension paradigm, slightly younger children could be tested than in Experiment 1.

Experiment 2

Method

Subjects. There were 36 adult subjects, all University of Illinois undergraduates, and 36 children: 18 4-5-year-olds, with a mean age of 4.10, and 18 7-8-year-olds, with a mean age of 7.11. They were approximately evenly divided between males and females.

Materials. Eighteen metaphors were used: eight R metaphors, eight D metaphors, and two A metaphors. The metaphors were taken from Experiments 1 (see Table 1), except that two of the D metaphors from Experiment 1 were judged unsatisfactory and replaced: the metaphors "A lake is like a mirror" and "Stars are like diamonds" were replaced by "Blood is like water" and "A submarine is like a whale." Because it proved to be difficult to design plausible relational interpretations for the A metaphors, only two A metaphors were used, both taken from Experiment 1: "Jelly beans are like balloons" and "A snake is like a hose." Due to experimenter error, the two new D metaphors were omitted in the adult version of the choice task, leaving only 16 metaphors: eight R, six D, and two A metaphors.

For each metaphor, two relational and two attributional interpretations were devised, based on the interpretations generated by subjects in Experiment 1. Three criteria were applied in selecting among the responses generated in Experiment 1. First, to be sure that the responses were either clearly relational or clearly attributional, the selection was limited to interpretations that had received scores of 4 or better from the judges in Experiment 1. Second, among the interpretations that met the first criterion, the two most frequent relational interpretations and the two most frequent attributional interpretations were selected. Third, because it was crucial that young children understand the choices, in tallying the responses those of 5-6-year-olds were given the most weight. For a few metaphors, there were not enough responses from Experiment 1 that met the criteria. In this case, I either used a less frequent response or invented an extra response to bring the total up to two attributional and two relational interpretations. Responses were reworded to simplify any difficult language. Table 3 shows samples of the metaphors and interpretations used.

[Insert Table 3 about here.]

Design and Procedure

In both the choice task and the rating task, the key factors were age (3 levels), metaphor type (R or D, with A considered separately), and interpretation type (relational or attributional). For adults, the tasks were administered in written form to two separate groups of 18 subjects. For the choice task, each metaphor was given with its two interpretations and subjects indicated which interpretation they preferred. For the rating task, each metaphor was given on a page with its two interpretations, and subjects were asked to rate how well they liked each interpretation on a 1-5 scale, with 5 being "very interesting" and 1 being "boring." In both tasks, the list of metaphors was presented twice, once with each set of interpretation pairs. Two different orders were used, one the reverse of the other.

Children were tested individually in their homes. They received the choice task followed by the rating task. A quiz-show format was used to simplify the task. Two stuffed dolls were used as quiz show contestants. Each doll had a canister to collect its winnings. The child was told that Pink Panther and Tweetie Bird were competing on a quiz show, and that it was the child's job to indicate which doll had given the best answer by dropping a token into the winning doll's canister. To be sure children understood the format, each child completed two practice tasks before starting the actual experiment.

Practice tasks. The first practice task was picture labeling. The experimenter held up a picture of a familiar animal and explained that the goal of the contestants was to give its common name (e.g., "elephant"). One doll produced the correct answer and the other a clearly incorrect answer. The child

indicated which doll was correct by dropping a token in that doll's canister. In this and all succeeding tasks, each doll was correct half the time, in a semirandom predetermined order. The second practice task was sentence interpretation. A simple sentence was read and each doll gave an interpretation of the meaning of the sentence. One interpretation was always correct and the other clearly wrong. For example, the sentence might be "Santa Claus is on the roof" and the two interpretations: "There's a jolly fat man on the roof" (correct), and "I had spaghetti for dinner" (wrong). The child indicated the correct answer in the manner described above. The criterion for passing the practice tasks was six correct in no more than 10 trials. In fact, most children achieved criterion in fewer than eight trials. Only one child--a 4-year-old--had to be dropped from the experiment due to failing the practice task. After the practice tasks, the experiment began with the choice task for metaphor interpretations.

Choice task. The experimenter read the "quiz question" metaphor: for example, "How is a cloud like a sponge?" Then the experimenter would point to each doll in turn and give the doll's response: "He says, 'Because they're both soft and fluffy'; and *he* says, 'They both hold water.' Whose answer is better?" The child would give a token to the winning doll. In all cases, one doll gave an attributional interpretation and the other a relational interpretation, with the assignment varying such that each doll had equal numbers of relational and attributional answers. Order of the relational and attributional interpretation was varied so that each occurred first half the time. After all 18 metaphors had been presented with the first set of interpretations, they were presented again with the second set of interpretations. The order in which the relational and attributional interpretations was given on the second pass was reversed from the order on the first pass.

Rating task. The experimenter began by saying, "Now we're going to do something else. You picked which answer you liked better. Now you're going to tell me just how good you think each answer is." Then the child was shown the "goodness meter," a vertical scale standing about 2 feet high and 3 inches wide, with gradations running from 1 (lowest) to 5 (highest). The experimenter pointed to the numbers and described their use: "A 5 means that the answer is really good. But a 1 means that it's really bad. You didn't like the answer at all. Now 3 is in the middle. It means that the answer was OK." When the child could demonstrate knowledge of the scale, the rating task began. The experimenter read the "quiz question" metaphor (e.g., "How is a cloud like a sponge?") and gave the first doll's response: "He says, 'Because they're both soft and fluffy.' How good is that answer?" After the child rated the answer on the goodness meter, the experimenter repeated the metaphor and gave the second doll's response, which the child again rated.

Definition task. At the end of the session, as a check on their knowledge, children were asked to define the words used in the metaphors. For younger children, all words were tested. For older children, only more difficult words (e.g., "submarine") were tested. If children had difficulty with the initial questions (e.g., "What is a sponge?"), the question was changed to "What do you do with a sponge?" All children demonstrated adequate understanding of the words.

Results and Discussion

Choice task. Relational responding increased with age, as shown in Figure 2.

[Insert Figure 2 about here.]

The proportions of relational responses for the R and D metaphors were .61, .69, and .89 for the 4-5-year-olds, 7-8-year-olds, and adults, respectively. A 3 (Age) x 2 (Metaphor Type) mixed-measures analysis of variance on the proportion of relational responses for the R and D metaphors showed a main effect of age, confirming the developmental increase in relational preference, $F(2,51) = 10.51, p < .0001$. (However, all age groups showed above-chance relational responding [Kolgoroff-Smirnov goodness-of-fit with chance at .50, all z's $> 1.50, p < .05$].) There was also a main effect of metaphor type, reflecting a stronger relational preference on R metaphors ($M = .77$) than on D metaphors ($M = .67$), $F(1,51) = 12.57, p < .001$. The interaction between age and metaphor type was not significant.

Rating task. There was an age increase in relational bias on the R and D metaphors. As shown in Figure 3, the mean rating of the relational interpretation increases with age, and the mean rating of the attributional interpretation decreases with age. These patterns were confirmed in a 3 (Age) x 2

Metaphor Type) x 2 (Interpretation Type) mixed-measures analysis of variance over the R and D metaphors. As predicted, there was a main effect of interpretation type, $F(1,51) = 95.88, p < .0001$, confirming that ratings were higher for relational interpretations than attributional interpretations. There was also a main effect of age, reflecting higher overall ratings for the 7-8-year-olds than for the other two groups, $F(2,51) = 6.27, p < .01$. The expected interaction of Age x Interpretation Type was significant, $F(2,51) = 30.83, p < .0001$, reflecting both the rise in ratings for the relational interpretations and the drop in ratings for the attributional interpretations with age. There was also a Metaphor x Interpretation Interaction, $F(1,51) = 13.94, p < .001$, indicating a somewhat greater relational preference on the R metaphors than on the D metaphors. There was no main effect of metaphor type and no other significant interactions.

[Insert Figure 3 about here.]

As shown in Figure 3, relational interpretations become steadily more valued with age and attributional interpretations become less valued. Planned comparisons confirmed that the older groups rated relational interpretations significantly higher than attributional interpretations. For R metaphors, the difference was significant for both 7-8-year-olds and adults, but not for the 4-5-year-olds, $t(17) = 5.50$, $p < .0001$; $t(17) = 10.17, p < .0001$; and $t(17) = 1.33, p > .05$, respectively. For D metaphors, only the adults showed a significant difference, $t(17) = 9.59, p < .0001$; neither the 4-5-year-olds nor the 7-8-year-olds showed a significant relational preference on D metaphors, $t(17) < 1$, and $t(17) = 1.89, p < .08$, respectively.

Attributional metaphors. On the A metaphors, subjects showed a clear preference for the attributional responses, as expected. In the choice task, all groups chose the attributional response more often than the relational response on the two A metaphors (see Fig. 2). Similarly, in the rating task all age groups rated the attributional response higher than the relational response on the two A metaphors (see Fig. 3). This difference was significant for 7-8-year-olds and adults, $t(17) = -2.71, p < .05$, and $t(17) = -4.91, p < .0001$, respectively, although not for 4-5-year-olds, $t(17) = -1.81, p < .09$. Performance on the A metaphors helps rule out some artifactual explanations of the results. The apparent relational preference on R and D metaphors might have been based on some response bias, such as a preference for longer[5] or more abstract-sounding interpretations. However, the subjects' preference for attributional interpretations on the A metaphors tends to indicate that their preference for relational interpretations on the R and D metaphors was a genuine interpretational choice.

General Discussion

The results of the two experiments show a developmental increase in relational focus in the interpretation of metaphors. As in prior research (Gentner, 1986), adults showed a strong relational bias. In Experiment 1, they generated relational interpretations whenever possible, and they rated metaphors as more apt the more relational commonalities they found and less apt the more attributional commonalities they found. In Experiment 2, on both the choice task and the rating task, adults showed a clear preference for relational interpretations for relational and double metaphors.

This relational focus developed gradually. In Experiment 1, the relationality of the interpretations produced for relational and double metaphors increased steadily from 6 years to adulthood. Experiment 2 showed the same developmental pattern for comprehension. Both on the choice task and on the ratings task, there was a developmental shift toward preferring relational interpretations over attributional interpretations of relational and double metaphors. In contrast, on the attributional metaphors, all three age groups preferred attributional interpretations both in production and in comprehension; there was no developmental change whatsoever. This indicates (a) that the developmental shift is specifically relational and not simply the result of some general increase in verbiage with age, and (b) that not all metaphors are alike developmentally. Indeed, young children seem to deal with attribute matches in much the same way as adults. What develops is relational similarity.

The notion of a relational shift helps clarify the conflicting findings in the development of metaphor. For example, the paradox of production before comprehension may admit of the following resolution: children both produce and comprehend attributional metaphors very early, but relational matches come

later. Because comprehension tasks typically include both kinds of metaphors, children do poorly in these tasks well after they can spontaneously produce attributional comparisons. A supporting observation is that most of the early spontaneous metaphors appear to be based on perceptual attribute overlap, as in Bowerman's "moon-hangnail" example quoted above. In Winner's (1979) study of the spontaneous metaphors of one child (Adam) from age 2.3 to 4.10, the most frequent basis for metaphorical comparison was the shared object attribute of *contour*: for example, referring to a pencil as a "big needle." Relational metaphors occasionally occurred, especially during pretending games (e.g., at 3.3 Adam placed a small alphabet letter on top of a big one and said, "Adam sleeping on Daddy"), but they were rare.

There is some prior research supporting a relational shift. In the study by Billow (1975) discussed above, metaphors based on shared attributes (e.g., "He had a pickle for a nose") were comprehended well before "proportional metaphors," based on shared relations (e.g., "The sun wakes the seeds"). The relational shift is also compatible with Keil's proposed shift from characteristic to defining features in category judgments (Keil & Batterman, 1984). Typically, characteristic features are object attributes, such as the shape of an animal, and defining features are relations, such as the causal regularity that parents produce children of the same species as themselves. Another antecedent is Bruner's proposal that children move from relying on perceptual-configural information to relying on functional information (Bruner, Olver, & Greenfield, 1966).

The notion of a relational shift allows a more orderly account of the development of metaphor. But what underlies this shift? At least three possibilities are suggested by the literature. The first possibility is that it reflects an increase in basic cognitive competence (Inhelder & Piaget, 1958; Pascual-Leone, 1970). Thus, Billow (1975) interpreted his findings in terms of cognitive stages: similar (attributional) metaphors are early because they can be understood at the concrete operational level; proportional (relational) metaphors are acquired later because they require formal operational thinking. The second possibility is that the relational shift reflects children's learning of adult pragmatic conventions concerning what to map. If children did not share the aesthetic that metaphors are supposed to be about shared relations, then their failure to perform relational mappings would not necessarily indicate a lack of competence. The third possibility is that the relational shift results from accretion of domain knowledge; it is an instance of the novice-expert shift (Carey, 1984; Chi, Feltovich, & Glaser, 1981; Larkin, 1983; Siegler, 1976). By this account, children's failure to perform relational mappings results not from cognitive deficiencies but from a lack of knowledge of the necessary domain relations.

Although there is as yet no definitive evidence, there is one line of comparison that can differentiate among these three approaches. The first two accounts--the cognitive competence account and the pragmatic insight account--both postulate a global shift that applies across domains. In contrast, the domain-knowledge account postulates that the relational shift will occur at different times in different domains. Moreover, it should occur earliest in those domains that are most accessible to young children. This position is supported by the fact that the few studies that have shown early ability to perform relational mappings have used extremely familiar domains (Brown & Campione, 1985; Crisafi & Brown, 1986; Gardner, 1974; Gentner, 1977a, 1977b). For example, in an earlier study I found that very young children could interpret metaphors with body parts as the base domain (Gentner, 1977a, 1977b). Children were shown pictures of familiar objects such as trees and asked questions like, "If the tree had a knee where would it be?" Children as young as 4 could accurately map the familiar spatial relations among body parts, even when the task was made more difficult by turning the tree upside down or by adding misleading details. Similarly, Crisafi and Brown (1986) found that preschoolers could learn to do a difficult combination-of-inferences problem by analogy with an easier problem, but only if the base problem was made up of extremely familiar objects and activities.

By the domain-knowledge account, the relational shift should occur early for some domains and late for others. DeLoache's (1985) work on search behavior provides intriguing evidence for a very early relational shift in spatial matching ability. At 3.1, children can solve problems in which they see an object hidden in a furnished room and then look for a tiny replica of the object in a tiny replica of the room. But at 2.7 children fail this task; they can successfully find the object only if the two rooms are identical. Perhaps it is during this interval that understanding of spatial relations progresses to the point where children can match on the basis of relational structure without the support of object identity. At the other end of the developmental spectrum, Sternberg and Downing (1982) found that adolescents go

through some of the same stages of interpretation when dealing with complex analogies between analogies that younger children do with simple analogies.

The domain-knowledge account of the relational shift still faces unanswered questions. In order to account for the data, the domain-knowledge account must presuppose that in every domain relational information is slower to be acquired than object-attribute information. This may be a plausible intuition, but it has not been verified. A second problem is specifying what constitutes a domain. Despite these unanswered questions, the domain-knowledge account is the best current candidate for explaining the relational shift.

References

Asch, S. E., & Nerlove, H. (1960). The development of double function terms in children: An exploratory investigation. In B. Kaplan & S. Wapner (Eds.), *Perspectives in psychological theory* (pp. 47-60). New York: International Universities Press.

Billow, R. M. (1975). A cognitive development study of metaphor comprehension. *Developmental Psychology, 11*, 415-423.

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 319-346). New York: Cambridge University Press.

Brown, A. L., & Campione, J. C. (1985). Three faces of transfer: Implications for early competence, individual differences, and instruction. In M. Lamb, A. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology* (Vol. 3, pp. 143-192). Hillsdale, NJ: Erlbaum.

Brunner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth*. New York: Wiley.

Carey, S. (1984). Are children fundamentally different kinds of thinkers and learners than adults? In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Current research and open questions* (Vol. 2, pp. 485-517). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-151.

Crisafi, M. A., & Brown, A. L. (1986). Analogical transfer in very young children: Combining two separately learned solutions to reach a goal. *Child Development, 57*, 573-576.

DeLoache, J. S. (1985). Memory-based searching in very young children. In H. Wellman (Ed.), *The development of search ability* (pp. 151-183). Hillsdale, NJ: Erlbaum.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1986). The structure-mapping engine. In *Proceedings of the American Association for Artificial Intelligence*. Philadelphia: American Association for Artificial Intelligence.

Gardner, H. (1974). Metaphors and modalities: How children project polar adjectives onto diverse domains. *Child Development, 45*, 84-91.

Gardner, H., Kircher, M., Winner, E., & Perkins, D. (1975). Children's metaphoric productions and preference. *Journal of Child Language, 2*, 1-17.

Gentner, D. (1977a). Children's performance on a spatial analogies task. *Child Development, 48*, 1034-1039.

Gentner, D. (1977b). If a tree had a knee, where would it be? Children's performance on simple spatial metaphors. *Papers and Reports on Child Language Development, 13*, 157-164.

Gentner, D. (1980). *The structure of analogical models in science*. (Tech. Rep. No. 4451). Cambridge, MA: Bolt Beranek & Newman.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D. (1986). *Evidence for a structure-mapping theory of analogy* (Tech. Rep.). Urbana-Champaign: University of Illinois.

Gentner, D., & Stuart, P. (1983). *Metaphor as structure-mapping: What develops.* (Tech. Rep. No. 5479). Cambridge, MA: Bolt Beranek & Newman.

Gentner, D., Falkenhainer, B., & Skorstad, J. (1987). Metaphor: The good, the bad and the ugly. In *Proceedings of the Third Conference on Theoretical Issues in Natural Language Processing,* Las Cruces, NM.

Gentner, D., & Landers, R. (1985). Analogical reminding: A good match is hard to find. In *Proceedings of the International Conference on Systems, Men, and Cybernetics* (pp. 607-613). Tucson, AZ.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10,* 277-300.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

Keil, F., & Batterman, N. A. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior, 23,* 221-236.

Kogan, N. (1975). *Metaphoric thinking in children: Developmental and individual-difference aspects.* Paper presented at the biennial meeting of the Society for Research in Child Development, Denver, CO.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Larkin, J. H. (1983). Problem representations in physics. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 75-98). Hillsdale, NJ: Erlbaum.

Miller, G. A. (1979). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (pp. 202-250). Cambridge: Cambridge University Press.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32,* 310-345.

Pollio, M. R., & Pickens, J. D. (1980). The developmental structure of figurative competence. In R. P. Honeck & R. R. Hoffman (Eds.), *Cognition and figurative language* (pp. 311-341). Hillsdale, NJ: Erlbaum.

Reynolds, R. E., & Ortony, A. (1980). Some issues in the measurement of children's comprehension of metaphorical language. *Child Development, 51,* 1110-1119.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481-520.

Sternberg, R. I., & Downing, C. J. (1982). The development of higher-order reasoning in adolescence. *Child Development, 53,* 209-221.

Thomson, J. R., & Chapman, R. S. (1975). Who is Daddy? The status of two-year-olds' overextended words in use and comprehension. *Papers and Reports on Child Language Development* (Stanford University), *10,* 59-68.

Verbrugge, R. R., & McCarrell, N. S. (1977). Metaphoric comprehension: Studies in reminding and resembling. *Cognitive Psychology, 9,* 494-533.

Vosniadou, S., & Ortony, A. (1983). The emergence of the literal-metaphorical-anomalous distinction in young children. *Child Development, 54,* 154-161.

Vosniadou, S., & Ortony, A. (1986). Testing the metaphoric competence of the young child: Paraphrase versus enactment. *Human Development, 29*, 226-230.

Winner, E. (1979). New names for old things: The emergence of metaphoric language. *Journal of Child Language, 6*, 469-491.

Footnotes

[1]The version of structure mapping presented here is a descriptive theory. A computer simulation based on a process model also exists (Falkenhainer, Forbus, & Gentner, 1986; Gentner, Falkenhainer, & Skorstad, 1987).

[2]For present purposes, metaphor ("An X is a Y") and simile ("An X is like a Y") will be classed together.

[3]This experiment was conducted by P. V. Stuart as an undergraduate honors project at UCSD.

[4]Although most scoring decisions were straightforward, there were some ambiguous cases. These generally involved relational adjectives such as *edible* and *soporific* or relational nouns such as *father*. These terms express relational information, yet they appear on the surface as adjectives (Miller, 1979). For example, the adjectival phrase "X is soporific" is on the surface a one-place predicate (soporific [X]). But its meaning is the relational proposition "X puts people to sleep." Because of their ambiguous status, such relational adjectives were scored as intermediate (3) on both the relational and the attributional scale.

[5]The relational responses were in general longer (in number of words) than the attributional responses, so a length bias would have inflated the number of relational responses. As another check on this possibility, I compared the results for four R and D metaphors that had the reverse pattern--attributional responses longer than relational responses--with the overall results for R and D metaphors. There were no systematic differences in the results.

Table 1

Materials Used in Experiment 1

Metaphor type	Examples
Relational metaphors	The moon is like a lightbulb. A camera is like a tape recorder. A ladder is like a hill. A cloud is like a sponge. A roof is like a hat. Treebark is like skin. A tire is like a shoe. A window is like an eye.
Attributive metaphors	Jelly beans are like balloons. A cloud is like a marshmallow. A football is like an egg. The sun is like an orange. A snake is like a hose. Soapsuds are like whipped cream. Pancakes are like nickels. A tiger is like a zebra.
Double metaphors	A doctor is like a repairman. A kite is like a bird. The sky is like the ocean. A hummingbird is like a helicopter. Plant stems are like drinking straws. A lake is like a mirror. Grass is like hair. Stars are like diamonds.

Table 2

Sample Interpretations of Different Classes of Metaphors in Experiment 1

Age	Attributional Metaphor (The sun is an orange.)	Relational Metaphor (A tire is a shoe.)	Double Metaphor (Plant stems are drinking straws.)
5-6 . . .	They're round and orange. (1,5)	Sometimes your shoe is black and the tire is black. (1,5)	They're both straight. (1,5)
9-10 . . .	It's like a circle and so is the sun. (1,5)	You can go places on both. (5,1)	Plant stems are thin and so are straws. (1,5)
Adult . . .	Both are orange; both spherical. (1,5)	Both are coverings that come in direct contact with the terrain. (5,1)	They are both used for drawing in water. They are both tubular. (5,5)

Note. The figures in parentheses give the rated relationality and attributionality of the responses.

Table 3

Sample Metaphors and Interpretations Used in Experiment 2

Samples	Interpretations	
Relational metaphors:		
A cloud is like a sponge	R1) Both can hold water	A1) Both are soft
	R2) Both can give off water	A2) Both are fluffy
A tire is like a shoe	R1) Both are used to move something	A1) Both are made of rubber
	R2) Both cover the bottom of something	A2) Both can be black
Double metaphors:		
Plant stems are like drinking straws. .	R1) Both can be used to get water	A1) Both are long
	R2) Both draw liquid up	A2) Both are thin
Grass is like hair .	R1) Both grow quickly	A1) Both are long
	R2) Both cover and protect something	A2) Both are skinny
Attributional metaphors:		
Jelly beans are like balloons	R1) Both are fun at parties	A1) Both can be different colors
	R2) Both can be tossed in the air	A2) Both are round
A snake is like a hose	R1) Both squirt liquid	A1) Both are long
	R2) Both can curl up	A2) Both are wiggley

Note. R1 and R2 denote relational interpretations, and A1 and A2 denote attributional interpretations.

Figure Captions

Figure 1a. Results of Experiment 1: Mean ratings of relationality for interpretations of different types of metaphor across age.

Figure 1b. Mean ratings of attributionality for interpretations of different types of metaphors across age.

Figure 2. Results of Experiment 2: Proportion of relational interpretations chosen for different types of metaphors across age.

Figure 3. Results of Experiment 2: Mean ratings of relational and attributional interpretations for different types of metaphors across age.

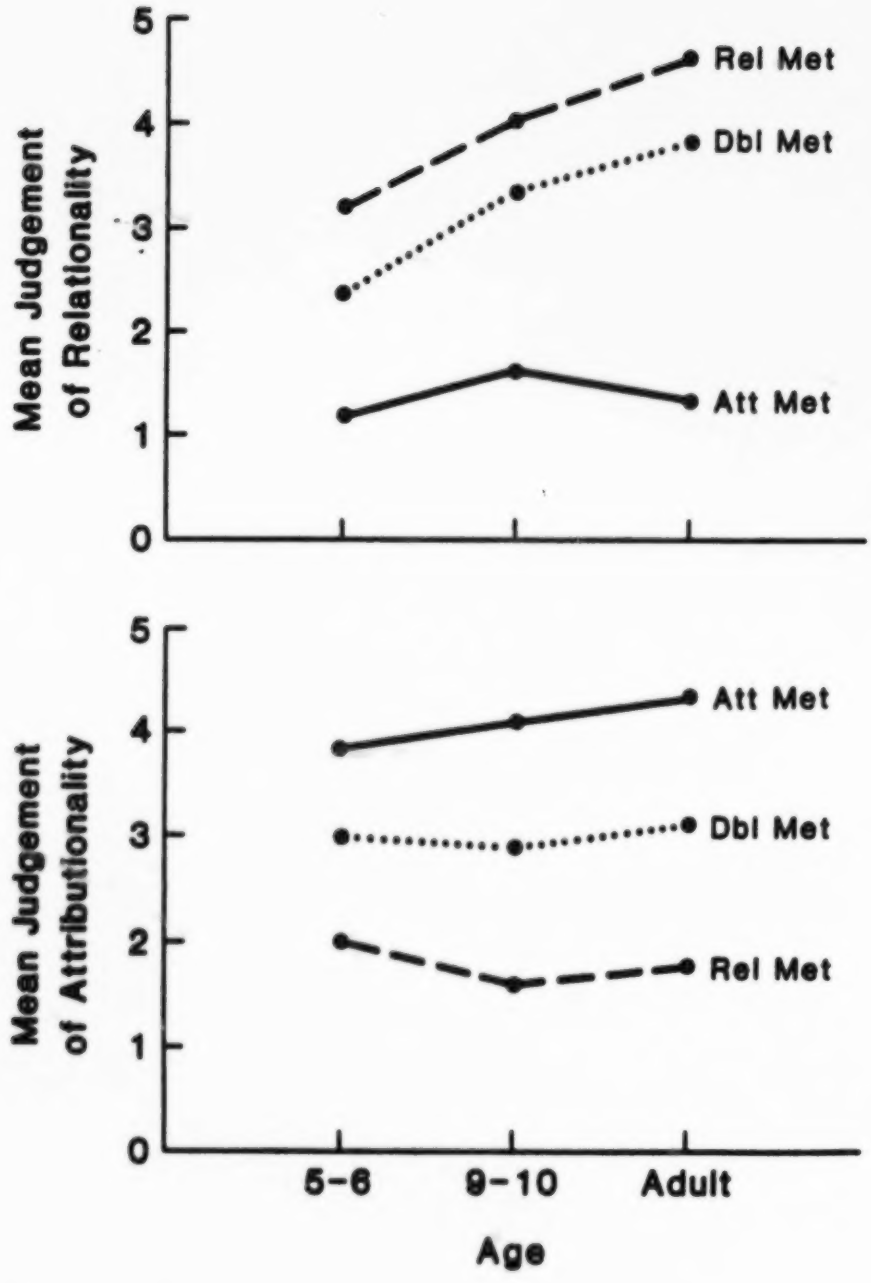
Mean Judgement of Relationality
5
4
3
2
1
0
Rel Met
Dbl Met
Att Met
Mean Judgement of Attributionality
5
4
3
2
1
0
Att Met
Dbl Met
Rel Met
5-6
9-10
Adult
Age

Figure 1a/1b

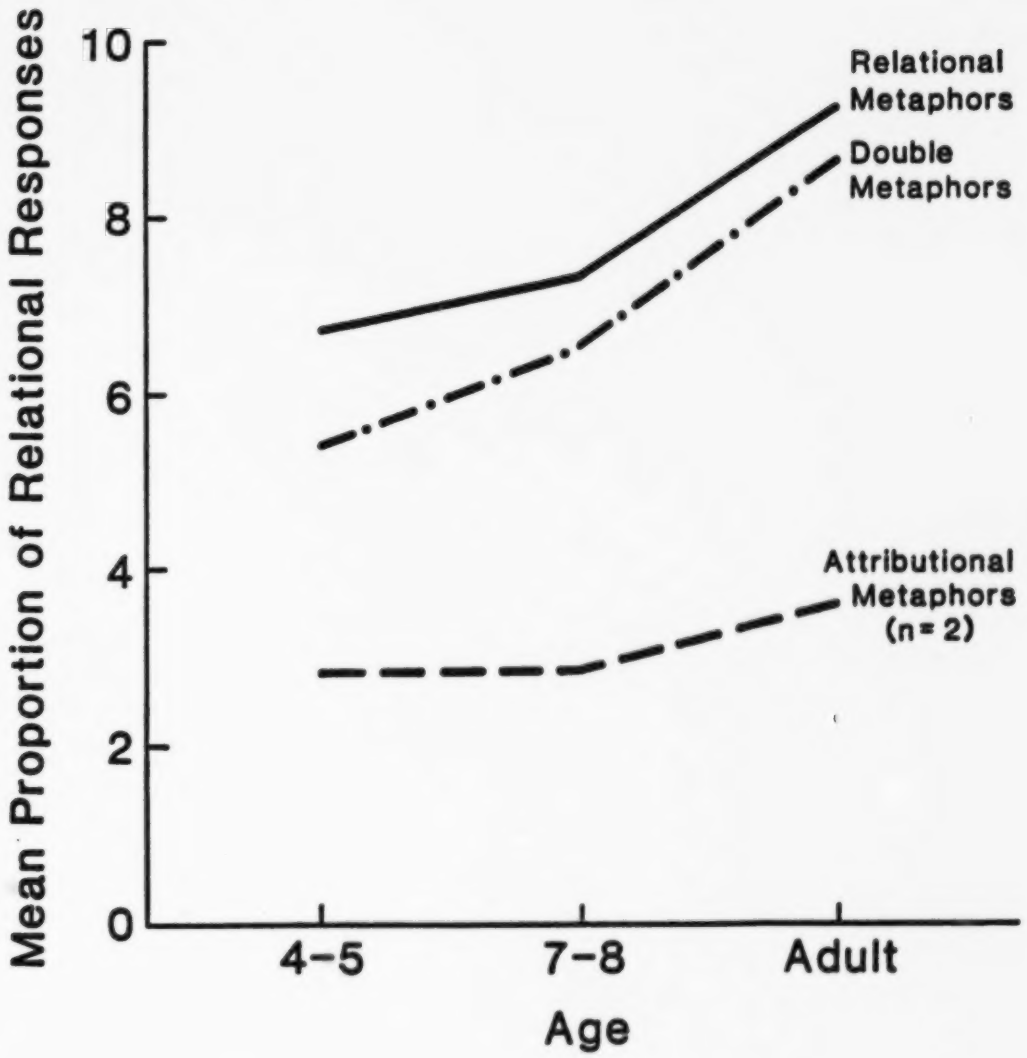
Mean Proportion of Relational Responses
10
8
6
4
2
0
Relational Metaphors
Double Metaphors
Attributional Metaphors (n = 2)
4-5
7-8
Adult
Age

Figure 2

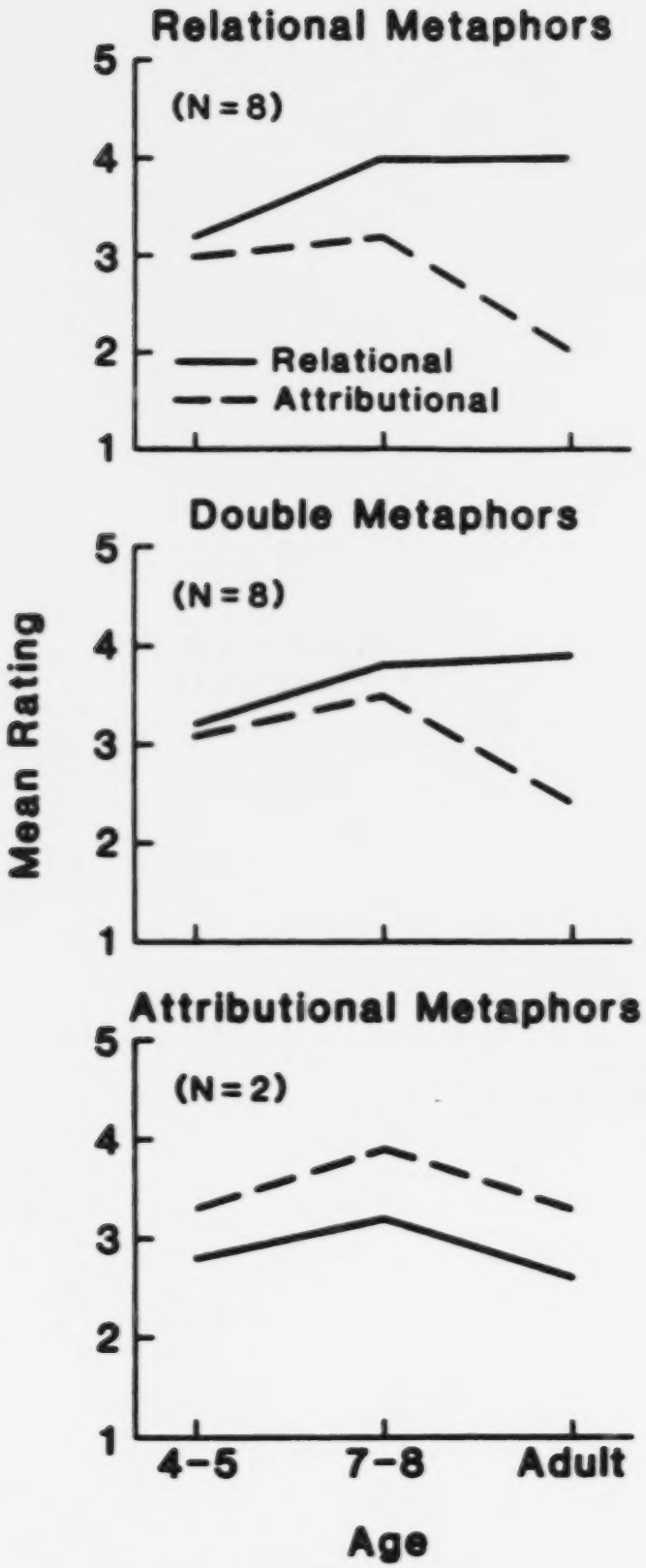
Relational Metaphors
(N = 8)
Relational
Attributional
Double Metaphors
(N = 8)
Attributional Metaphors
(N = 2)
Mean Rating
5
4
3
2
1
4–5
7–8
Adult
Age

Figure 3